



LATHAM & WATKINS LLP

MORRISON FÖRSTER

November 21, 2024

Hon. Ona T. Wang
Daniel Patrick Moynihan
United States Courthouse
500 Pearl St.
New York, NY 10007-1312

cc: All Counsel of Record Line (via ECF)

Re: *Authors Guild et al. v. OpenAI, Inc., et al.*, Case Nos. 1:23-cv-08292-SHS & 1:23-cv-10211-SHS [[ECF 262 – Plaintiffs’ Letter Motion Seeking Search Terms for RFP No. 98](#)]

Dear Magistrate Judge Wang:

For the reasons set forth below, Plaintiffs’ request to compel OpenAI to run a hit report for fourteen *non*-custodians should be denied. Plaintiffs’ request is merely an attempt to improperly circumvent the parties’ stipulated ESI Order. Further, it is disproportionate and unnecessary, given the extensive discovery that OpenAI has already agreed to provide in response to RFP No. 98 and other similar RFPs.¹

I. Plaintiffs’ request improperly circumvents the custodial discovery process.

Plaintiffs’ request that OpenAI run numerous search terms across fourteen former employees’ ESI—when none of these fourteen individuals are custodians—plainly violates the ESI Order. The Court should deny Plaintiffs’ request on at least this basis.

Custodial discovery is governed by the Stipulated ESI Order, which was carefully negotiated by the parties and entered by the Court on September 13, 2024. *See* ECF 201. The ESI Order specifies that hit reports are required solely *for custodians*. *See id.* ¶ 4. Hit reports are not required in other circumstances. The reason is obvious: hit reports are generated by running search terms across data that has been collected from a custodian, processed to be searchable, and stored in a database—a series of steps that is expensive and time-consuming for each custodian. *See, e.g., Nike, Inc. v. Wu*, No. 13-CIV-8012 (CM), 2020 WL 257475, at *12 (S.D.N.Y. Jan. 17, 2020) (“Negotiating search terms, identifying email custodians, and weeding out non-responsive emails are some of the most time-consuming stages of 21st-century document discovery.”).

OpenAI has already designated twenty-four custodians in this case—four of whom are specifically named in RFP No. 98: Dario Amodei, Greg Brockman, Ben Mann, and John Schulman. *See* ECF 262 at n.1, 2. For each of the existing twenty-four custodians, OpenAI has shouldered the significant burden of collecting, processing, storing, searching, and reviewing

¹ Although Plaintiffs’ letter refers to broad “safety related concerns,” (ECF 262 at 2), that has nothing to do with the subject of RFP No. 98, which seeks documents related to concerns about use of “copyrighted material” and the “general harm or risk to content creators.”

their data. OpenAI has agreed to run more than 600 search terms across the custodial data. Ex. A ¶ 2. OpenAI’s custodial review already encompasses more than 460,000 documents.² See ECF 255 at 2.

In recognition of the ESI process, Plaintiffs initially proposed that the dispute over RFP No. 98 could be resolved by running search terms on *existing* custodians’ files. See Ex. B at 5 (memorializing a meet and confer in which Plaintiffs indicated they “would agree to propose search terms to run on the existing custodians’ files to identify the information they seek in response to this RFP”). After all, four of the existing custodians are specifically named in RFP No. 98. ECF 262 at n.1, 2. OpenAI has agreed to apply Plaintiffs’ requested search terms to the existing custodians. But Plaintiffs have reneged on their proposed approach, instead pivoting to demand that OpenAI run search terms across fourteen additional former OpenAI employees who are *not* custodians—in essence, attempting to get fourteen additional custodians through the back door. Ex. B at 2-3. This is improper, and endorsing Plaintiffs’ approach would eviscerate the custodial process agreed to by the parties and endorsed by the Court.

II. OpenAI is providing ample discovery on RFP No. 98 and related topics, rendering Plaintiffs’ request disproportionate and unnecessary.

Even assuming Plaintiffs’ request complied with the ESI Order (it does not), Plaintiffs’ request is disproportionate and unnecessary given OpenAI’s discovery efforts pertaining to RFP No. 98 and similar requests. RFP No. 98 seeks documents relating to “statements made or concerns raised” by twenty-three individuals regarding (a) OpenAI’s “use of purportedly copyrighted material in training Large Language Models;” (b) OpenAI’s “lack of or insufficient internal controls with regard to acquisition and use of training materials or control over output;” and (c) “the general harm or risk to content creators posed by” OpenAI’s LLMs.³ ECF 262-1 at 6. OpenAI’s discovery efforts, spanning custodial and non-custodial repositories, amply address this RFP.

First, with respect to RFP No. 98 in particular, OpenAI has agreed to run Plaintiffs’ proposed search terms for existing custodians, including across the custodial data for four individuals named in RFP No. 98. See Ex. B at 1 (agreeing to run Plaintiffs’ proposed terms for RFP No. 98 on existing custodians’ files). Plaintiffs recognize there is no dispute on this point. See ECF 262 at 2 n.2.

Second, OpenAI has agreed to identify and search non-custodial repositories for complaints related to alleged copyright infringement—the precise issue targeted by RFP No. 98. OpenAI’s search has encompassed a review of all submissions to OpenAI’s DMCA/Copyright Disputes form, and all submissions to an employee hotline. Ex. C at 1.

² Meanwhile, just last month, Plaintiffs renewed a motion to add eight additional custodians (ECF 230).

³ Underscoring the disproportionate nature of Plaintiffs’ request, five of the individuals named in RFP No. 98 are unaffiliated with OpenAI whatsoever. See ECF 262 at 2 n.5 (identifying the five individuals who are unaffiliated with OpenAI: Kumar, Nanda, Benigo, Hinton, Russell).

Third, OpenAI has agreed to produce documents in response to requests covering similar topics as RFP No. 98. For example, OpenAI has agreed to produce documents in response to RFPs that address “complaints regarding [OpenAI’s] use of purportedly copyrighted material in training” (RFP No. 7); “any efforts to mitigate and/or remedy instances in which [OpenAI] purportedly used copyrighted material without permission in training ChatGPT” (RFP No. 27); and “the bases of [OpenAI’s] aware[ness] that creators are concerned about being replaced or displaced by AI” (RFP No. 96). *See* Ex. D (OpenAI’s responses to RFP Nos. 7, 8, 9); Ex. E (OpenAI’s responses to RFP Nos. 25, 27, 28); Ex. F (OpenAI’s responses to RFP Nos. 95 and 96).

In sum, given OpenAI’s efforts, it is disproportionate and unnecessary to require OpenAI to collect, process, store, and search the custodial data from fourteen additional individuals for purposes of responding to RFP No. 98. OpenAI is already identifying and searching the repositories that likely contain responsive information relating to this RFP.⁴ There is no reason to believe that these fourteen additional individuals will have more unique information that is responsive. Indeed, Plaintiffs rely solely on these individuals’ comments regarding *general* safety—which is distinct from the three specific categories at issue in RFP No. 98. *See, e.g.*, ECF 262 at 2 (“A Right to Warn about Advanced Artificial Intelligence” discussing only “serious risks posed by these technologies”); *id.* at 2 n.6 (Aschenbrenner’s comments about OpenAI’s general safety practices). Plaintiffs have not justified their request.

* * *

Plaintiffs’ request reflects an improper attempt to evade custodial discovery, and to impose scorch-the-earth discovery on OpenAI. The Court should deny Plaintiffs’ request.

Respectfully,

KEKER, VAN NEST &
PETERS LLP

LATHAM & WATKINS LLP

MORRISON & FOERSTER
LLP

/s/ Katie Lynn Joyce
Katie Lynn Joyce

/s/ Elana Nightingale Dawson
Elana Nightingale Dawson

/s/ Carolyn M. Homer
Carolyn M. Homer

⁴ “[OpenAI] as the responding party is best situated to decide how to search for and produce ESI responsive to [Plaintiffs’] document requests.” *Hyles v. New York City*, No. 10-CIV-3119 (AT)(AJP), 2016 WL 4077114, at *3 (S.D.N.Y. Aug. 1, 2016).